Uniformly More Powerful Tests for Hypotheses Concerning Linear Inequalities and Normal Means

ROGER L. BERGER*

This article considers some hypothesis-testing problems regarding normal means. In these problems, the hypotheses are defined by linear inequalities on the means. We show that in certain problems the likelihood ratio test (LRT) is not very powerful. We describe a test that has the same size, α, as the LRT and is uniformly more powerful. The test is easily implemented, since its critical values are standard normal percentiles. The increase in power with the new test can be substantial. For example, the new test's power is $1/2\alpha$ times bigger (10 times bigger for $\alpha = .05$) than the LRT's power for some parameter points in a simple example.

Specifically, let $\mathbf{X} = (X_1, \ldots, X_p)'$ $(p \geq 2)$ be a multivariate normal random vector with unknown mean $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$ and known, nonsingular covariance matrix $\boldsymbol{\Sigma}$. We consider testing the null hypothesis H_0: $\mathbf{b}_i'\boldsymbol{\mu} \leq 0$ for some $i = 1, \ldots, k$ versus the alternative hypothesis H_1: $\mathbf{b}_i'\boldsymbol{\mu} > 0$ for all $i = 1, \ldots, k$. Here $\mathbf{b}_1, \ldots, \mathbf{b}_k$ $(k \geq 2)$ are specified p-dimensional vectors that define the hypotheses. Many types of relationships among the means may be described with the linear inequalities. Two interesting types are those that specify the signs of the means and those that describe an order relationship. Some examples of alternative hypotheses that can be specified in this way are these: H_1^s: $\mu_i > 0$, $i = 1, \ldots, p$ (sign testing), H_1^o: $\mu_1 < \mu_2 < \cdots < \mu_p$ (simple order), H_1^L: $\mu_1 < \mu_i < \mu_p$, $i = 2, \ldots, p - 1$ (simple loop), and H_1^T: $\mu_1 < \mu_i$, $i = 2, \ldots, p$ (simple tree). If $\mu_i = \nu_{2i} - \nu_{1i}$, where ν_{ji} is the average response of the ith patient subset to the jth treatment, then H_1^s states that Treatment 2 is better than Treatment 1 for all subsets. If the μ_i are regression coefficients, then H_1^s states that the mean response increases with each independent variable. In any case, these relationships would be the alternative hypothesis. Rejection of H_0 by a test with small size would be taken as strong evidence confirming that the specified sign or order relationship is true.

Sasabuchi (1980) showed that the size-α LRT of H_0 versus H_1 is the test that rejects H_0 if $Z_i = \mathbf{b}_i'\mathbf{X}/(\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_i)^{1/2} \geq z_\alpha$ for all $i = 1, \ldots, k$, where z_α is the upper 100α percentile of a standard normal distribution. This test is biased and has very low power if all of the values $\mathbf{b}_i'\boldsymbol{\mu}$ $(i = 1, \ldots, k)$ are only slightly bigger than 0. We define an integer J and constants $c_0, \ldots, c_J$ that are certain standard normal percentiles. We show that, in many cases, a size-α test that is uniformly more powerful than the LRT is the test that rejects H_0 if $\mathbf{X} \in R_1 \cup \cdots \cup R_J$, where $R_j = \{\mathbf{x}: c_j \leq z_i \leq c_{j-1}, i = 1, \ldots, k\}$ and $z_i = \mathbf{b}_i'\mathbf{x}/(\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_i)^{1/2}$ is the LRT statistic. The set R_1 is the rejection region of the LRT, so this test is obviously more powerful than the LRT. But we show that if, for each $i = 1, \ldots, k$, there exists an $m \neq i$ such that $\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_m \leq 0$, then this test is also a size-α test. It is easy to verify that this condition is satisfied, for example, for all of the aforementioned H_1 hypotheses, except the simple tree, if $\boldsymbol{\Sigma}$ is diagonal.

Tests that are even more powerful than those just described might exist. We discuss an example of such a test. But despite this test's superior power properties, it has some counterintuitive properties. Thus tests such as in this example may be primarily of theoretical interest.

All of the previously mentioned results are derived in the $\boldsymbol{\Sigma}$-known case. Sasabuchi (1980) showed that, if $\boldsymbol{\Sigma}$ is unknown, the LRT is very similar. The differences are that $\boldsymbol{\Sigma}$ is replaced by an estimate and z_α is replaced by t_α, a t-distribution percentile. We show, in an example, that making the same modifications to this test does not give a size-α test. But in the example the size of the test converges to α quickly as the degrees of freedom for the estimate of $\boldsymbol{\Sigma}$ becomes large. So even for moderate degrees of freedom ($\geq$10), this test might be preferable to the LRT, since its size is approximately α and it is much more powerful than the LRT.

A two-sided version of this problem is obtained if we test H_0^2: $\boldsymbol{\mu} \notin (H_1 \cup -H_1)$ versus H_1^2: $\boldsymbol{\mu} \in (H_1 \cup -H_1)$, where H_1 is a one-sided alternative as defined above. Sasabuchi (1980) showed that the LRT rejects H_0^2 if $Z_i \geq c$ for all $i = 1, \ldots, k$ or $Z_i \leq -c$ for all $i = 1, \ldots, k$, Sasabuchi gave some conditions under which $c = z_\alpha$ gives a size-α test. We consider only the special case in which H_1 is the sign-testing alternative and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2)$, a diagonal matrix. For constants $c_0, \ldots, c_{2J}$, similar to those above, we show that the test that rejects H_0^2 if $\mathbf{X} \in R_1 \cup \cdots \cup R_{2J}$, where $R_j = \{\mathbf{x}: c_j \leq x_i/\sigma_i \leq c_{j-1}, i = 1, \ldots, p\}$, is a size-$\alpha$ test that is uniformly more powerful than the LRT. For the special case of $p = 2$, this provides a test that is uniformly more powerful than a test proposed by Gail and Simon (1985) for testing for a qualitative interaction.

KEY WORDS: Likelihood ratio test; Majorization; Polyhedral cone; Qualitative interaction.

1. TESTING PROBLEM AND LIKELIHOOD RATIO TEST

Let $\mathbf{X}' = (X_1, \ldots, X_p)$ be a p-variate $(p \geq 2)$ normal random vector with unknown mean $\boldsymbol{\mu}$ and nonsingular covariance matrix $\boldsymbol{\Sigma}$. We will use the notation $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Throughout the article, except in Section 5, $\boldsymbol{\Sigma}$ will be assumed known. The results in this article can be considered approximately true if $\boldsymbol{\Sigma}$ is unknown but a large sample is available for estimating $\boldsymbol{\Sigma}$. In many applications, $\boldsymbol{\Sigma}$ will be a diagonal matrix; that is, the p populations with means $\mu_1, \ldots, \mu_p$ will be independent populations and X_i will be the sample mean of a random sample from the ith population. But we will consider the more general setting.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ be k $(k \geq 2)$ specified p-dimensional vectors. We consider testing

the null hypothesis

$$H_0\text{: } \mathbf{b}_i'\boldsymbol{\mu} \leq 0 \quad \text{for some } i = 1, \ldots, k$$

* Roger L. Berger is Associate Professor, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203. This research was funded in part by Subgrant 652 from the U.S. Army Research Office through the Academy of Applied Science. Kristin Rahn and Soku Yi are thanked for computer-programming assistance that produced tables. Authur Cohen is thanked for helpful discussions. The associate editor and two referees are thanked for helpful suggestions.

Journal of the American Statistical Association
March 1989, Vol. 84, No. 405, Theory and Methods

versus the alternative hypothesis

$$H_1: \mathbf{b}_i'\boldsymbol{\mu} > 0 \quad \text{for all } i = 1, \ldots, k. \tag{1.1}$$

For this to be meaningful, H_1 must be nonempty. (We use the symbol H_1 to denote the set of $\boldsymbol{\mu}$ vectors specified by the hypothesis, as well as the statement of the hypothesis.) This would not be the case, for example, if $\mathbf{b}_1 = -\mathbf{b}_2$. We assume that there are no redundant vectors in $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$. That is, there is no $\mathbf{b}_j$ such that $\{\boldsymbol{\mu}: \mathbf{b}_i'\boldsymbol{\mu} > 0, i = 1, \ldots, k\} = \{\boldsymbol{\mu}: \mathbf{b}_i'\boldsymbol{\mu} > 0, i = 1, \ldots, k, i \neq j\}$. This requirement only simplifies notation and proofs and in no way restricts the testing problems we are considering. Sasabuchi (1980) discussed conditions that are equivalent to the requirement that H_1 is nonempty and $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ has no redundant vectors.

Sasabuchi (1980) showed that the size-α likelihood ratio test (LRT) of H_0 versus H_1 is the test that rejects H_0 if

$$Z_i = \mathbf{b}_i'\mathbf{X}/(\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_i)^{1/2} \geq z_\alpha \quad \text{for all } i = 1, \ldots, k, \tag{1.2}$$

where z_α is the upper 100α percentile of the standard normal distribution. Berger (1982) and Cohen, Gatsonis, and Marden (1983a) discussed some applications of this test.

Actually, Sasabuchi (1980) considered a slightly different testing problem. The null hypothesis considered by Sasabuchi was

$$H_0: \mathbf{b}_i'\boldsymbol{\mu} \geq 0 \quad \text{for all } i = 1, \ldots, k,$$

with equality for at least one i. Sasabuchi's alternative hypothesis was the same as ours. In some cases our formulation may be more appropriate because the hypotheses do not artificially restrict the natural parameter space of $\boldsymbol{\mu}$. It is easy to modify Sasabuchi's argument to see that, in either case, the LRT has a rejection region of the following form: Reject H_0 if $Z_i \geq c$ for all $i = 1, \ldots, k$. To show that (1.2) is the size-α LRT for (1.1), it remains to show that $c = z_\alpha$ yields a size-α test. Our null hypothesis is a much larger set than Sasabuchi's. So when we take the supremum over H_0 of the rejection probability, we could get a larger size. But, in fact, the suprema over both sets are the same (see Sec. 3) and (1.2) does define the size-α LRT in our problem.

The LRT has two optimality properties. Lehmann (1952) and Cohen et al. (1983b) proved that the LRT is uniformly most powerful among all monotone, level-α tests. The more powerful tests we describe are not monotone. Cohen et al. (1983b) also showed that, in a bivariate problem, the LRT is *admissible* in that no other test has a uniformly smaller power function on H_0 and a uniformly bigger power function on H_1. Nomakuchi and Sakata (1987) generalized this result. The more powerful tests we describe dominate the LRT in that they have the same size, α, and uniformly bigger power on H_1. But they do not dominate the LRT in this decision-theoretic sense.

Despite these good properties the LRT has some deficiencies. It is a biased test. The power is less than α for some $\boldsymbol{\mu} \in H_1$. In fact, the bias can be quite extreme. For example, suppose $\boldsymbol{\Sigma} = \text{diag}\,(\sigma_1^2, \ldots, \sigma_p^2)$ is a diagonal matrix and consider the sign-testing problem

$$H_0^S: \mu_i \leq 0 \quad \text{for some } i = 1, \ldots, p$$

versus

$$H_1^S: \mu_i > 0 \quad \text{for all } i = 1, \ldots, p\,. \tag{1.3}$$

The LRT rejects H_0^S if $Z_i = X_i/\sigma_i \geq z_\alpha$ for all $i = 1, \ldots, p$. If $\boldsymbol{\mu} = \mathbf{0}$, $Z_1, \ldots, Z_p$ are independent $N_1(0, 1)$ random variables. So the power at $\boldsymbol{\mu} = \mathbf{0}$ is $P_{\mathbf{0}}(Z_1 \geq z_\alpha, \ldots, Z_p \geq z_\alpha) = \alpha^p$, which is much less than α. Of course $\boldsymbol{\mu} = \mathbf{0} \in H_0$, but the power function is continuous. So for $\boldsymbol{\mu} \in H_1$ that are close to $\mathbf{0}$, the power will be approximately α^p. To some extent this bias is unavoidable. Lehmann (1952) showed that in some problems of this type, *no* unbiased, nonrandomized test exists. Nomakuchi and Sakata (1987) also discussed this. But tests do exist that have the same size as the LRT and are uniformly more powerful. Tests with this property are described in Sections 3 and 4. For the aforementioned problem, the test in Section 3 has power equal to $\alpha^{p-1}/2$ at $\boldsymbol{\mu} = \mathbf{0}$. Thus this test's power is $(\alpha^{p-1}/2)/\alpha^p = 1/2\alpha$ times as big as the LRT's at some parameter points. This is a tenfold increase if $\alpha = .05$ and a fiftyfold increase if $\alpha = .01$.

Tests that are uniformly more powerful than the LRT are not unknown. Gutmann (1987) demonstrated the existence of a test that was uniformly more powerful than the uniformly most powerful monotone test in the sign-testing problem (1.3) when $X_1, \ldots, X_p$ are independent. Gutmann considered a general location model. In the normal problem, our tests in Sections 3 and 4 are uniformly more powerful than Gutmann's test and hence provide an affirmation of the conjecture made by Gutmann in his example (Gutmann 1987, p. 283). Warrack and Robertson (1984) and Berger and Sinclair (1984) described other problems in which the LRT can be dominated.

Robertson and Wegman (1978) found the LRT for the testing problem in which H_1 is the null hypothesis and H_0 is the alternative hypothesis. That is, the null hypothesis states that $\boldsymbol{\mu}$ is in a cone and the alternative hypothesis states that $\boldsymbol{\mu}$ is not in the cone. The test statistic is quite different, involving isotonic regression estimates of $\boldsymbol{\mu}$, and the critical values are percentiles for weighted sums of chi-squared or beta distributions.

The type of inference one wishes to make and the error one wishes to guard against determine whether the Robertson and Wegman formulation or the formulation in (1.1) is appropriate. For example, suppose $\mu_1, \ldots, \mu_p$ are regression coefficients and the sign-testing hypothesis H_1^S from (1.3) is suggested by a theory. If the experimenter only wished to abandon the theory if the data strongly suggest that it was false, then H_1^S should be the null hypotheses. This is a goodness-of-fit–type situation. But if the experimenter wanted to know if the data provided strong evidence confirming the theory, then H_1^S should be the alternative, as in (1.3). Rejection of H_0^S by a test with small α would be strong evidence that all of the inequalities in H_1^S are true. Sometimes one formulation is appropriate and sometimes the other is. But failure to reject the null hypothesis H_1^S, in the Robertson and Wegman formulation, cannot be taken as strong confirmation that all of the inequalities in H_1^S are true.

For many computations, it is more convenient to consider this transformed version of the original problem that was used by Sasabuchi (1980). Let $\mathbf{T}$ be a $p \times p$ nonsingular matrix such that $\mathbf{T\Sigma T}' = \mathbf{I}_p$, the $p \times p$ identity matrix. Thus $\mathbf{T}^{-1}(\mathbf{T}^{-1})' = \mathbf{\Sigma}$. Make the transformation $\mathbf{Y} = \mathbf{TX}$. Then $\mathbf{Y} \sim N_p(\boldsymbol{\theta}, \mathbf{I}_p)$, where $\boldsymbol{\theta} = \mathbf{T}\boldsymbol{\mu}$. Define $\mathbf{a}_1, \ldots, \mathbf{a}_k$ by $\mathbf{a}_i' = \mathbf{b}_i'\mathbf{T}^{-1}$. Then $\mathbf{b}_i'\boldsymbol{\mu} = \mathbf{a}_i'\boldsymbol{\theta}$. Thus our original testing problem is equivalent to observing $\mathbf{Y}$ and testing

$$H_0\colon \mathbf{a}_i'\boldsymbol{\theta} \le 0 \quad \text{for some } i = 1, \ldots, k$$

versus

$$H_1\colon \mathbf{a}_i'\boldsymbol{\theta} > 0 \quad \text{for all } i = 1, \ldots, k.$$

The LRT rejects H_0 if $Z_i = \mathbf{a}_i'\mathbf{Y}/(\mathbf{a}_i'\mathbf{a}_i)^{1/2} \ge z_\alpha$ for all $i = 1, \ldots, k$. We will consistently use the notation $\mathbf{X}$, $\boldsymbol{\mu}$, and $\mathbf{b}_i$ for quantities in the original problem and $\mathbf{Y}$, $\boldsymbol{\theta}$, and $\mathbf{a}_i$ for quantities in the transformed problem.

In Section 2, we prove some preliminary results that will be used to show that various tests are size-α tests. Readers may only wish to read the theorems' statements on first reading. But Definitions 2.1 and 2.2 should be noted. In Section 3 we describe a size-α test that is uniformly more powerful than the LRT. We compare the powers of the two tests for the sign-testing problem (1.3) when $p = 2$. In Section 4 we discuss an even more powerful test for the sign-testing problem (1.3). In Section 5, the sign-testing problem (1.3) with $p = 2$ is considered with an unknown variance. In Section 6, a two-sided version of the problem is considered and a size-α test that is uniformly more powerful than the LRT is described for a sign-testing problem.

2. PRELIMINARY THEOREMS

The following results will be used to prove that various tests are size-α tests. For any vector $\mathbf{g}$, define $\|\mathbf{g}\| = (\mathbf{g}'\mathbf{g})^{1/2}$.

Lemma 2.1. Let $\mathbf{g}$ and $\mathbf{h}$ be noncolinear vectors ($|\mathbf{g}'\mathbf{h}| < \|\mathbf{g}\|\,\|\mathbf{h}\|$) satisfying $\mathbf{g}'\mathbf{h} \le 0$. Let

$$\mathbf{d} = \left[\mathbf{h} - \left(\frac{\mathbf{g}'\mathbf{h}}{\mathbf{g}'\mathbf{g}}\right)\mathbf{g}\right] \Big/ \left\|\mathbf{h} - \left(\frac{\mathbf{g}'\mathbf{h}}{\mathbf{g}'\mathbf{g}}\right)\mathbf{g}\right\|$$

be the unique (up to sign) vector of length 1 in the space spanned by $\mathbf{g}$ and $\mathbf{h}$ that is orthogonal to $\mathbf{g}$ ($\mathbf{d}'\mathbf{g} = 0$) and let

$$r = [(\|\mathbf{g}\|\,\|\mathbf{h}\| - \mathbf{g}'\mathbf{h})/(\|\mathbf{g}\|\,\|\mathbf{h}\| + \mathbf{g}'\mathbf{h})]^{1/2}.$$

If $\mathbf{y}$ is a vector and c is a scalar such that $\mathbf{g}'\mathbf{y} \ge c\|\mathbf{g}\|$ and $\mathbf{h}'\mathbf{y} \ge c\|\mathbf{h}\|$, then $\mathbf{d}'\mathbf{y} \ge cr$. If $\mathbf{y}$ and c satisfy $\mathbf{g}'\mathbf{y} \le c\|\mathbf{g}\|$ and $\mathbf{h}'\mathbf{y} \le c\|\mathbf{h}\|$, then $\mathbf{d}'\mathbf{y} \le cr$.

Proof. We prove the first result. Replace $\mathbf{y}$ with $-\mathbf{y}$ to prove the second result. The conditions on $\mathbf{g}$ and $\mathbf{h}$ imply that $\mathbf{g} \ne 0$ and $\mathbf{h} \ne 0$; hence all ratios are well defined. Let δ and γ denote the coefficients on $\mathbf{h}$ and $\mathbf{g}$, respectively, in $\mathbf{d}$. Note that $\delta > 0$ and $\gamma \ge 0$. Hence $\mathbf{d}'\mathbf{y} = \delta(\mathbf{h}'\mathbf{y}) + \gamma(\mathbf{g}'\mathbf{y}) \ge \delta(c\|\mathbf{h}\|) + \gamma(c\|\mathbf{g}\|)$. Substituting the expressions for γ and δ and simplifying the expressions yields the result.

The constants $c_0, \ldots, c_{2J}$, used to define the rejection regions for our tests, are defined as follows.

Definition 2.1. For $0 < \alpha < .5$, define the integer J by the inequality $J - 1 < 1/2\alpha \le J$. Define the constants $c_0, \ldots, c_{2J}$ as follows: $c_0 = \infty$, $c_j = z_{j\alpha}$ ($j = 1, \ldots, J - 1$), $c_J = 0$, and $c_j = -c_{2J-j}$ ($j = J + 1, \ldots, 2J$).

Notice that $c_0 > c_1 > \cdots > c_{2J}$. If $1/2\alpha$ is an integer (as it is for $\alpha = .10$, $.05$, and $.01$), then $c_1, c_2, \ldots, c_{2J-1}$ are the $N_1(0, 1)$ percentiles, $z_\alpha, z_{2\alpha}, \ldots, z_{(2J-1)\alpha}$. For any α, if $Z \sim N_1(0, 1)$, then $\Pr(c_j \le Z \le c_{j-1}) = \alpha$ for $j = 1, \ldots, J - 1$ and $j = J + 2, \ldots, 2J$. $\Pr(c_J \le Z \le c_{J-1}) = \Pr(c_{J+1} \le Z \le c_J) \le \alpha$ with equality if $1/2\alpha$ is an integer.

Lemma 2.2. Let $\mathbf{g}$ and $\mathbf{h}$ satisfy the conditions in Lemma 2.1. Define the sets $S_1^*, \ldots, S_{2J}^*$ by

$$S_j^* = \left\{\mathbf{y}\colon c_j \le \frac{\mathbf{g}'\mathbf{y}}{\|\mathbf{g}\|} \le c_{j-1},\ c_j \le \frac{\mathbf{h}'\mathbf{y}}{\|\mathbf{h}\|} \le c_{j-1}\right\}. \tag{2.1}$$

Let $\mathbf{Y} \sim N_p(\boldsymbol{\theta}, \mathbf{I}_p)$. If $\mathbf{g}'\boldsymbol{\theta} = 0$, then $P_{\boldsymbol{\theta}}(\mathbf{Y} \in \bigcup_{j=1}^{2J} S_j^*) \le \alpha$.

Proof. Let $\mathbf{d}$ and r be as in Lemma 2.1. Define the sets $S_1^+, \ldots, S_{2J}^+$ by

$$S_j^+ = \left\{\mathbf{y}\colon c_j \le \frac{\mathbf{g}'\mathbf{y}}{\|\mathbf{g}\|} \le c_{j-1},\ c_j r \le \mathbf{d}'\mathbf{y} \le c_{j-1}r\right\}.$$

Lemma 2.1 implies that $S_j^* \subset S_j^+$. Also,

$$S_j^+ \cap S_{j+1}^+ \subset \{\mathbf{y}\colon \mathbf{g}'\mathbf{y}/\|\mathbf{g}\| = c_j\},$$

a set with probability 0, and $S_j^+ \cap S_i^+ = \phi$ if $|j - i| > 1$. Thus

$$P_{\boldsymbol{\theta}}\left(\mathbf{Y} \in \bigcup_{j=1}^{2J} S_j^*\right) \le P_{\boldsymbol{\theta}}\left(\mathbf{Y} \in \bigcup_{j=1}^{2J} S_j^+\right) = \sum_{j=1}^{2J} P_{\boldsymbol{\theta}}(\mathbf{Y} \in S_j^+). \tag{2.2}$$

The random variables $\mathbf{g}'\mathbf{Y}/\|\mathbf{g}\|$ and $\mathbf{d}'\mathbf{Y}$ are independent normal random variables, since $\mathbf{d}'\mathbf{I}_p\mathbf{g} = \mathbf{d}'\mathbf{g} = 0$. And $\mathbf{g}'\mathbf{Y}/\|\mathbf{g}\|$ has a standard normal distribution if $\mathbf{g}'\boldsymbol{\theta} = 0$. Thus

$$\begin{aligned}
&\sum_{j=1}^{2J} P_{\boldsymbol{\theta}}(\mathbf{Y} \in S_j^+) \\
&\quad = \sum_{j=1}^{2J} P_{\boldsymbol{\theta}}\left(c_j \le \frac{\mathbf{g}'\mathbf{Y}}{\|\mathbf{g}\|} \le c_{j-1},\ c_j r \le \mathbf{d}'\mathbf{Y} \le c_{j-1}r\right) \\
&\quad = \sum_{j=1}^{2J} P_{\boldsymbol{\theta}}\left(c_j \le \frac{\mathbf{g}'\mathbf{Y}}{\|\mathbf{g}\|} \le c_{j-1}\right) P_{\boldsymbol{\theta}}(c_j r \le \mathbf{d}'\mathbf{Y} \le c_{j-1}r) \\
&\qquad \text{(independence)} \\
&\quad \le \sum_{j=1}^{2J} \alpha P_{\boldsymbol{\theta}}(c_j r \le \mathbf{d}'\mathbf{Y} \le c_{j-1}r) \\
&\qquad \text{(property of } c_0, \ldots, c_{2J}) \\
&\quad = \alpha.
\end{aligned}$$

This with (2.2) yields the desired result.

The construction in the proof of Lemma 2.2 is illustrated in Figure 1 for the case when $p = k = 2$, $\alpha = .2$, $\mathbf{g}' = (-1, 2)$, and $\mathbf{h}' = (1, 0)$. In this case $1/2\alpha = 2.5$, so $J = 3$, $c_1 = -c_5 = .84$, $c_2 = -c_4 = .25$, and $c_3 = 0$. The diamond-shaped regions are the sets $S_1^*, \ldots, S_6^*$. The rectangular regions with dashed borders are the sets $S_1^+, \ldots, S_6^+$. Note that $S_j^* \subset S_j^+$. All of the edges of $S_1^*, \ldots, S_6^*$ are perpendicular to either $\mathbf{g}$ or $\mathbf{h}$. The angle between $\mathbf{g}$ and $\mathbf{h}$ is at least 90°; that is, $\mathbf{g}'\mathbf{h} \le 0$. The angle η—that is, 180° minus the angle between $\mathbf{g}$ and $\mathbf{h}$—is at most 90°. Because $\eta \le 90°$, we can construct the rectangles to contain the diamonds. The dashed lines with negative slope correspond to the $\mathbf{y}$'s for which $\mathbf{d}'\mathbf{y} = c_j r$ $(j = 1, \ldots, 5)$.

The rejection regions for the tests we will consider are formed from the following sets.

Definition 2.2. Let $z_i = z_i(\mathbf{x}) = \mathbf{b}_i'\mathbf{x}/(\mathbf{b}_i'\mathbf{\Sigma}\mathbf{b}_i)^{1/2}$. For α, J, and $c_0, \ldots, c_{2J}$ as in Definition 2.1, define the following sets:

$$R_j = \{\mathbf{x}\colon c_j \le z_i \le c_{j-1},\ i = 1, \ldots, k\}, \qquad j = 1, \ldots, 2J.$$

Under the transformation $\mathbf{y} = \mathbf{T}\mathbf{x}$ described in Section 1, the set R_j is mapped onto the set

$$S_j = \left\{\mathbf{y}\colon c_j \le \frac{\mathbf{a}_i'\mathbf{y}}{(\mathbf{a}_i'\mathbf{a}_i)^{1/2}} = \frac{\mathbf{a}_i'\mathbf{y}}{\|\mathbf{a}_i\|} \le c_{j-1},\ i = 1, \ldots, k\right\}.$$

The following theorems will be used to show that various tests are of size α. We state them in terms of the original quantities $\mathbf{X}$, $\boldsymbol{\mu}$, and $\mathbf{b}_i$, as that is the context in which they will be used.

Theorem 2.1. Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$. Suppose the set $(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ is such that H_1 in (1.1) is nonempty. Suppose further that for each $i = 1, \ldots, k$ there is an $m \in \{1, \ldots, k\}$ (m will depend on i) such that $\mathbf{b}_i'\mathbf{\Sigma}\mathbf{b}_m \le 0$. Let $0 < \alpha < .5$, and let $c_0, \ldots, c_{2J}$ and $R_1, \ldots, R_{2J}$ be as in Definitions 2.1 and 2.2. If $\boldsymbol{\mu}$ satisfies $\mathbf{b}_i'\boldsymbol{\mu} = 0$ for some $i \in \{1, \ldots, k\}$, then $P_{\boldsymbol{\mu}}(\mathbf{X} \in \bigcup_{j=1}^{2J} R_j) \le \alpha$.

Proof. Using the transformation $\mathbf{Y} = \mathbf{T}\mathbf{X}$, define $\boldsymbol{\theta} = \mathbf{T}\boldsymbol{\mu}$. Note that $\mathbf{a}_i'\boldsymbol{\theta} = \mathbf{b}_i'\mathbf{T}^{-1}\mathbf{T}\boldsymbol{\mu} = 0$. Let m be such that $\mathbf{b}_i'\mathbf{\Sigma}\mathbf{b}_m \le 0$. Then $\mathbf{a}_i'\mathbf{a}_m = \mathbf{b}_i'\mathbf{T}^{-1}(\mathbf{T}^{-1})'\mathbf{b}_m = \mathbf{b}_i'\mathbf{\Sigma}\mathbf{b}_m \le 0$. Since H_1 is nonempty, $\mathbf{a}_i$ and $\mathbf{a}_m$ are noncolinear ($\mathbf{a}_i'\mathbf{a}_m$ cannot be less than $-\|\mathbf{a}_i\| \cdot \|\mathbf{a}_m\|$, and $\mathbf{a}_i'\mathbf{a}_m = -\|\mathbf{a}_i\| \cdot \|\mathbf{a}_m\|$ implies that $\mathbf{a}_m = -f\mathbf{a}_i$ for some positive constant f; but this would imply that H_1 is empty). Thus $\mathbf{a}_i$ and $\mathbf{a}_m$ satisfy the conditions on $\mathbf{g}$ and $\mathbf{h}$ in Lemmas 2.1 and 2.2. Notice that with $\mathbf{g} = \mathbf{a}_i$ and $\mathbf{h} = \mathbf{a}_m$, S_j from Definition 2.2 is a subset of S_j^* from (2.1). Thus from Lemma 2.2 we have

$$P_{\boldsymbol{\mu}}\Big(\mathbf{X} \in \bigcup_{j=1}^{2J} R_j\Big) = P_{\boldsymbol{\theta}}\Big(\mathbf{Y} \in \bigcup_{j=1}^{2J} S_j\Big) \le P_{\boldsymbol{\theta}}\Big(\mathbf{Y} \in \bigcup_{j=1}^{2J} S_j^*\Big) \le \alpha.$$

The second theorem is quite general and unrelated to the special structure we have used up to now. But we have not found it stated in the literature in this generality.

Theorem 2.2. Let $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$. Let R be a set and $\mathbf{b}$ be a vector such that $\mathbf{b}'\mathbf{x} \ge 0$ for every $\mathbf{x} \in R$. Let $\boldsymbol{\mu}$ be a vector such that $\mathbf{b}'\boldsymbol{\mu} \le 0$. Then there exists a vector $\boldsymbol{\mu}^*$ such that $\mathbf{b}'\boldsymbol{\mu}^* = 0$ and $P_{\boldsymbol{\mu}^*}(\mathbf{X} \in R) \ge P_{\boldsymbol{\mu}}(\mathbf{X} \in R)$.

Proof. If $\mathbf{b}'\boldsymbol{\mu} = 0$ then $\boldsymbol{\mu}^* = \boldsymbol{\mu}$ satisfies the requirements. So assume that $\mathbf{b}'\boldsymbol{\mu} < 0$. Let $\mathbf{T}$ be the nonsingular matrix defined in Section 1. Let $\mathbf{a}' = \mathbf{b}'\mathbf{T}^{-1}$. Now let $\mathbf{O} = (\mathbf{o}_1, \ldots, \mathbf{o}_p)'$ be an orthogonal matrix with $\mathbf{o}_1 = \mathbf{a}/\|\mathbf{a}\|$. Make the transformation $\mathbf{U} = \mathbf{O}\mathbf{T}\mathbf{X}$. Let $P = \mathbf{O}\mathbf{T}R = \{\mathbf{u}\colon \mathbf{u} = \mathbf{O}\mathbf{T}\mathbf{x}, \mathbf{x} \in R\}$. Then $\mathbf{U} \sim \mathbf{N}_p(\boldsymbol{\nu} = \mathbf{O}\mathbf{T}\boldsymbol{\mu}, \mathbf{I}_p)$ and $P_{\boldsymbol{\nu}}(\mathbf{U} \in P) = P_{\boldsymbol{\mu}}(\mathbf{X} \in R)$. For every $\mathbf{u} \in P$, $u_1 = \mathbf{b}'\mathbf{x}/\|\mathbf{a}\| \ge 0$. Also, $\nu_1 = \mathbf{b}'\boldsymbol{\mu}/\|\mathbf{a}\| < 0$. Thus

$$\begin{aligned}
&P_{\boldsymbol{\nu}}(\mathbf{U} \in P)\\
&= \int \cdots_P \int (2\pi)^{-p/2} \exp -\left(\frac{1}{2}\sum_{i=1}^{p} (u_i - \nu_i)^2\right) du_1 \cdots du_p\\
&\le \int \cdots_P \int (2\pi)^{-p/2}\\
&\qquad \times \exp -\left(\frac{1}{2}(u_1 - 0)^2 + \frac{1}{2}\sum_{i=2}^{p} (u_i - \nu_i)^2\right) du_1 \cdots du_p\\
&= P_{\boldsymbol{\nu}^*}(\mathbf{U} \in P),
\end{aligned}$$

where $\boldsymbol{\nu}^{*\prime} = (0, \nu_2, \ldots, \nu_p)$. Now making the two inverse transformations we have $\boldsymbol{\mu}^* = \mathbf{T}^{-1}\mathbf{O}'\boldsymbol{\nu}^*$ and $P_{\boldsymbol{\mu}^*}(\mathbf{X} \in R) = P_{\boldsymbol{\nu}^*}(\mathbf{U} \in P) \ge P_{\boldsymbol{\mu}}(\mathbf{X} \in R)$. Furthermore, since $\nu_1^* = 0$ and $\mathbf{O}$ is orthogonal we have

$$\mathbf{b}'\boldsymbol{\mu}^* = \mathbf{b}'\mathbf{T}^{-1}\mathbf{O}'\boldsymbol{\nu}^* = \mathbf{a}'\left(\sum_{i=2}^{p} \nu_i\mathbf{o}_i\right) = \|\mathbf{a}\|\mathbf{o}_1'\left(\sum_{i=2}^{p} \nu_i\mathbf{o}_i\right) = 0.$$

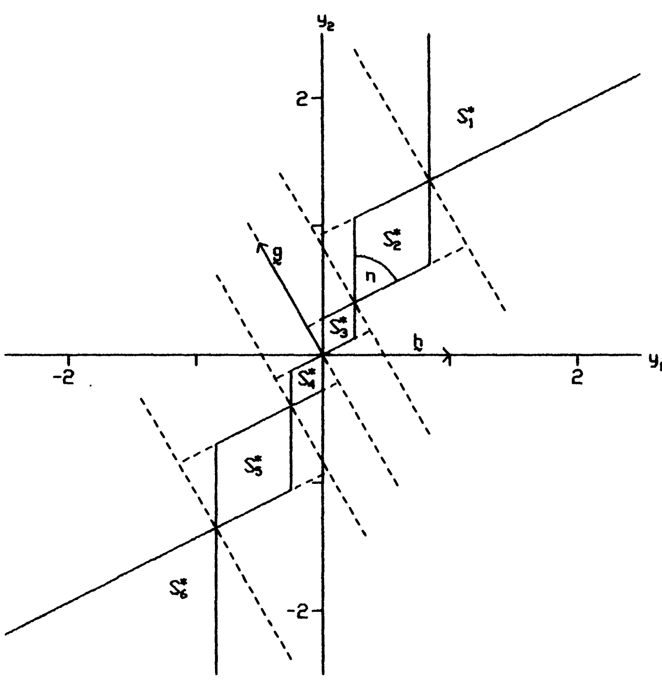

Figure 1. Sets From Lemma 2.2. The conditions on **g** *and* **h** *ensure that the diamond-shaped sets,* $S_1^*, \ldots, S_6^*$, *can be enclosed in the dashed rectangles.*

3. A TEST THAT IS MORE POWERFUL THAN THE LRT

Under certain conditions, the following test will be shown to be a size-α test that is uniformly more powerful than the LRT for the testing problem described in (1.1).

Definition 3.1. For values of α that satisfy $0 < \alpha < .5$, define Test I as the test that rejects H_0 if, for some $j \in (1, \ldots, J)$, $c_j \le Z_i \le c_{j-1}$ for all $i = 1, \ldots, k$, where $c_0, \ldots, c_J$ are from Definition 2.1 and $Z_i = \mathbf{b}_i'\mathbf{X}/(\mathbf{b}_i' \boldsymbol{\Sigma} \mathbf{b}_i)^{1/2}$. Alternatively, the rejection region for Test I can be expressed as $R_1 \cup \cdots \cup R_J$, where the sets R_j are from Definition 2.2.

Example 3.1. Let $p = k = 2$. Suppose X_1 and X_2 are independent and $X_i \sim N_1(\mu_i, \sigma_i^2)$. Let $\mathbf{b}_1' = (1, 0)$ and $\mathbf{b}_2' = (0, 1)$ so that we are testing (1.3). Then $Z_i = X_i/\sigma_i$. If $\alpha = .10$, then $J = 5$, $c_1 = 1.28$, $c_2 = .84$, $c_3 = .52$, $c_4 = .25$, and $c_5 = 0$. So Test I's rejection region consists of the five rectangles, $R_1 \cup \cdots \cup R_5$, in Figure 2. R_1 is the rejection region for the LRT.

Example 3.2. Let X_1 and X_2 be as in Example 3.1. Consider testing H_0: $2\mu_2 \le \mu_1$ or $\mu_1 \le 0$ versus H_1: $0 < \mu_1 < 2\mu_2$. Then $\mathbf{b}_1' = (-1, 2)$ and $\mathbf{b}_2' = (1, 0)$. If $\sigma_1 = \sigma_2 = 1$ and $\alpha = .2$, the rejection region for Test I is $S_1^* \cup S_2^* \cup S_3^*$ in Figure 1, where the axes are now the $x_1 - x_2$ axes. For a smaller, more commonly used, value of α, the picture would be similar but with more (but smaller) diamond-shaped regions.

We now prove that Test I has the properties we desire.

Theorem 3.1. For the testing problem described in (1.1), suppose that for each $i = 1, \ldots, k$ there exists an $m \in \{1, \ldots, k\}$ (m will depend on i) such that $\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_m \le 0$. If $0 < \alpha < .5$, then Test I is a size-α test and Test I is uniformly more powerful than the size-α LRT.

Proof. The size-α LRT, as found by Sasabuchi (1980), rejects H_0 if $Z_i \ge z_\alpha$ for all $i = 1, \ldots, k$. But $c_0 = \infty$ and $c_1 = z_\alpha$. So R_1 is the rejection region of the size-α LRT. Since R_1 is a subset of the rejection region Test I, Test I is uniformly more powerful than the size-α LRT.

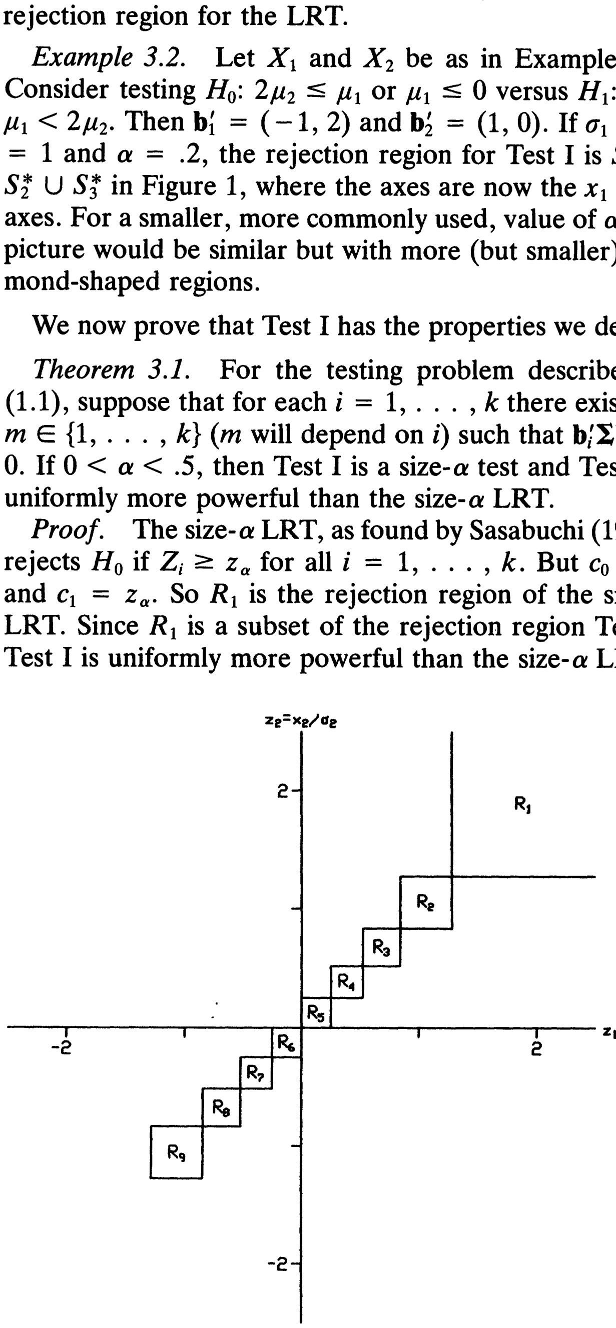

Figure 2. Rejection Regions for LRT, Test I, and Test II in the Bivariate Sign-Testing Problem With $\alpha = .10$. The rejection region for LRT is R_1. The rejection region for Test I is $R_1 \cup \cdots \cup R_5$. The rejection region for Test II is $R_1 \cup \cdots \cup R_9$.

Let $H_s = (\boldsymbol{\mu}\colon \mathbf{b}_i'\boldsymbol{\mu} \ge 0$ for all $i = 1, \ldots, k$ and $\mathbf{b}_i'\boldsymbol{\mu} = 0$ for some $i)$. Sasabuchi showed that $\sup_{\boldsymbol{\mu}\in H_s} P_{\boldsymbol{\mu}}(\mathbf{X} \in R_1) = \alpha$; that is, the LRT is a size-α test for the null hypothesis, H_s. But $H_s \subset H_0$; hence

$$\alpha = \sup_{\boldsymbol{\mu}\in H_s} P_{\boldsymbol{\mu}}(\mathbf{X} \in R_1) \le \sup_{\boldsymbol{\mu}\in H_0} P_{\boldsymbol{\mu}}(\mathbf{X} \in \bigcup_{j=1}^{J} R_j)$$

$$= \text{size of Test I.} \tag{3.1}$$

Now let $\boldsymbol{\mu} \in H_0$. Then there exists an i such that $\mathbf{b}_i'\boldsymbol{\mu} \le 0$. For all $\mathbf{x} \in R_1 \cup \cdots \cup R_J$, $\mathbf{b}_i'\mathbf{x}/(\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_i)^{1/2} = z_i \ge c_J = 0$; hence $\mathbf{b}_i'\mathbf{x} \ge 0$. By Theorem 2.2, there is a $\boldsymbol{\mu}^*$ with $\mathbf{b}_i'\boldsymbol{\mu}^* = 0$ such that

$$P_{\boldsymbol{\mu}^*}(\mathbf{X} \in \bigcup_{j=1}^{J} R_j) \ge P_{\boldsymbol{\mu}}(\mathbf{X} \in \bigcup_{j=1}^{J} R_j). \tag{3.2}$$

By Theorem 2.1, the conditions on $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ imply that

$$\alpha \ge P_{\boldsymbol{\mu}^*}(\mathbf{X} \in \bigcup_{j=1}^{2J} R_j) > P_{\boldsymbol{\mu}^*}(\mathbf{X} \in \bigcup_{j=1}^{J} R_j). \tag{3.3}$$

Since $\boldsymbol{\mu} \in H_0$ was arbitrary, (3.1), (3.2), and (3.3) imply that Test I is a size-α test.

It may seem curious that one can take a size-α test (the LRT), add sets of positive probability to the rejection region, and still have a size-α test. Although $\sup_{\boldsymbol{\mu}\in H_s} P_{\boldsymbol{\mu}}(\mathbf{X} \in R_1) = \alpha$, this is possible because $P_{\boldsymbol{\mu}}(\mathbf{X} \in R_1) < \alpha$ for every $\boldsymbol{\mu} \in H_0$. Sasabuchi (1980) showed that the supremum was only attained in a limit as one $\mathbf{b}_i'\boldsymbol{\mu} = 0$ and all other $\mathbf{b}_i'\boldsymbol{\mu} \to \infty$. For all $\boldsymbol{\mu} \in H_0$, Test I's power function satisfies $P_{\boldsymbol{\mu}}(\mathbf{X} \in R_1) < P_{\boldsymbol{\mu}}(\mathbf{X} \in \bigcup_{j=1}^{J} R_j) < \alpha$.

The restriction in Theorem 3.1, that for each i there exists an m such that $\mathbf{b}_i'\boldsymbol{\Sigma}\mathbf{b}_m \le 0$, is a restriction on the hypothesis-testing problems for which we have shown that Test I is a more powerful size-α test than the LRT. If $\boldsymbol{\Sigma} = \mathbf{I}_p$, then this restriction, in light of Lemma 2.1 with $c = 0$, has the following geometric interpretation. For each subspace $\mathbf{b}_i'\boldsymbol{\mu} = 0$ that contains a face of the polyhedral cone H_1, there is another subspace, defined by $\mathbf{d}'\boldsymbol{\mu} = 0$, such that the two subspaces are perpendicular, in the sense that $\mathbf{b}_i'\mathbf{d} = 0$, and the cone H_1 lies entirely between these two subspaces. So the restriction says that the cone cannot be too spread out. Here are four examples of alternative hypotheses H_1:

Sign testing
H_1^S: $\mu_i > 0$ $(i = 1, \ldots, p)$.
Simple order
H_1^O: $\mu_1 < \mu_2 < \cdots < \mu_p$.
Simple loop
H_1^L: $\mu_1 < \mu_i < \mu_p$ $(i = 2, \ldots, p-1)$.
Simple tree
H_1^T: $\mu_1 < \mu_i$ $(i = 2, \ldots, p)$.

All but the simple tree satisfy this condition. This definition of perpendicular subspaces is not the usual definition of orthogonal subspaces. But it is what one would mean in three dimensions if one thought of two planes being perpendicular. Two two-dimensional planes cannot

Table 1. Power for LRT, Test I, and Test II in the Bivariate Sign-Testing Problem With $\alpha = .10$

	μ					
	0	.5	1	2	3	4
$\beta_L(0, \mu)$	.010	.022	.039	.076	.096	.100
$\beta_I(0, \mu)$	.050	.069	.084	.098	.100	.100
$\beta_{II}(0, \mu)$	.090	.096	.099	.100	.100	.100
$\beta_L(\mu, \mu)$	.010	.047	.151	.583	.916	.993
$\beta_I(\mu, \mu)$	.050	.105	.209	.600	.917	.993
$\beta_{II}(\mu, \mu)$	.090	.124	.215	.600	.917	.993
$\beta_L(.5\mu, \mu)$	.010	.033	.085	.297	.561	.761
$\beta_I(.5\mu, \mu)$	.050	.087	.141	.327	.567	.762
$\beta_{II}(.5\mu, \mu)$	.090	.110	.152	.328	.567	.762

be orthogonal in three dimensions according to the usual definition.

To illustrate quantitatively the improvement in power that Test I provides, consider again Example 3.1. We use $\sigma_1 = \sigma_2 = 1$ and $\alpha = .10$. The rejection region for Test I is $R_1 \cup \cdots \cup R_5$ in Figure 2, and the rejection region for the LRT is just R_1. Let $\beta_I(\boldsymbol{\mu})$ and $\beta_L(\boldsymbol{\mu})$ be the power functions of Test I and the LRT, respectively. Values of these two functions for certain $\boldsymbol{\mu}$ values are in Table 1. [The third value, $\beta_{II}(\boldsymbol{\mu})$, is the power of a test from Sec. 4.] The top part of the table is for values of $\boldsymbol{\mu}' = (0, \mu)$, $\mu \geq 0$. These values are on the boundary of H_0, so the power is everywhere less than $\alpha = .10$. An unbiased test would have power equal to $\alpha = .10$ for all of these $\boldsymbol{\mu}$ values. Test I and the LRT are biased, but Test I is less so than the LRT. The power comparison mentioned after (1.3) can be made for this example: $\beta_I(\mathbf{0})/\beta_L(\mathbf{0}) = .05/.01 = 5 = 1/2\alpha$. In the middle of Table 1 are values of the power for mean vectors on the diagonal, $\boldsymbol{\mu}' = (\mu, \mu)$, $\mu \geq 0$. $\beta_I(\boldsymbol{\mu})$ is noticeably above $\beta_L(\boldsymbol{\mu})$ for $\mu \leq 2$ with the largest difference, $\beta_I(\boldsymbol{\mu}) - \beta_L(\boldsymbol{\mu}) \approx .07$, occurring in the range $.5 < \mu < 1$. The bottom of Table 1 contains values of the power function for mean vectors of the form $\boldsymbol{\mu}' = (.5\mu, \mu)$, $\mu \geq 0$. $\beta_I(\boldsymbol{\mu})$ is noticeably larger than $\beta_L(\boldsymbol{\mu})$ for $\mu \leq 3$ with the maximum difference, $\beta_I(\boldsymbol{\mu}) - \beta_L(\boldsymbol{\mu}) \approx .06$, occurring in the range $.5 < \mu < 1.1$.

4. AN EVEN MORE POWERFUL TEST

Test I is not necessarily the most powerful size-α test. In some cases there exist size-α tests that are uniformly more powerful than Test I. In this section we give an example of such a test, Test II.

Test II will reject H_0 if $\mathbf{X} \in R_1 \cup \cdots \cup R_M$, where $J < M < 2J$. The rejection region for Test II consists of the rejection region for Test I plus more of the sets R_j. Test II is obviously more powerful than Test I or the LRT. But the verification that Test II is a size-α test is more difficult. Theorem 2.2 cannot be used because the rejection region does not lie on one side of a plane.

Test II may be primarily of theoretical interest because it has a rather counterintuitive property. For any $\mathbf{x} \in R_j$ $(j > J)$, $\mathbf{b}_i'\mathbf{x} \leq 0$ for all $i = 1, \ldots, k$. Thus, if we reject H_0 for such an $\mathbf{x}$, we are deciding that $\mathbf{b}_i'\boldsymbol{\mu} > 0$ for all $i = 1, \ldots, k$ even though $\mathbf{x}$, the estimate of $\boldsymbol{\mu}$, satisfies $\mathbf{b}_i'\mathbf{x} \leq 0$ for all $i = 1, \ldots, k$. Although M can be chosen so that Test II is a size-α test that is uniformly more powerful than Test I, the important question might be this: Is there a size-α test with power comparable to Test II that only rejects for $\mathbf{x}$ such that $\mathbf{b}_i'\mathbf{x} \geq 0$ for all $i = 1, \ldots, k$?

Again consider Example 3.1: $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \mathbf{I}_2)$ (for simplicity we set both variances equal to 1). For $\alpha = .10$, Test I has the rejection region $R_1 \cup \cdots \cup R_5$, where the R_j are in Figure 2. We will show that Test II, with rejection region $R_1 \cup \cdots \cup R_9$ also has size $\alpha = .10$. To compute the size of Test II we use majorization techniques. See Marshall and Olkin (1974) for all definitions regarding these concepts. Each of the sets R_j is a Schur-convex set, and any union of Schur-convex sets is a Schur-convex set. Thus the rejection region for Test II, for any M, is a Schur-convex set. The density of $\mathbf{X}$ is Schur concave. By theorem 2.1 of Marshall and Olkin (1974), the power function of Test II, $\beta_{II}(\boldsymbol{\mu})$, is a Schur-concave function. That is, if $\boldsymbol{\mu}$ majorizes $\boldsymbol{\mu}^*$, then $\beta_{II}(\boldsymbol{\mu}^*) \geq \beta_{II}(\boldsymbol{\mu})$.

The size of Test II is $\sup_{\boldsymbol{\mu} \in H_0} \beta_{II}(\boldsymbol{\mu})$. We wish to determine the largest $M > J$ (if any exists) for which the size is α. Let $\boldsymbol{\mu} \in H_0$ with $\mu_1 + \mu_2 \geq 0$. Then $\boldsymbol{\mu}$ majorizes $\boldsymbol{\mu}^* = (\mu_1 + \mu_2, 0)'$. The rejection region of Test II, for any M, is a subset of $R_1 \cup \cdots \cup R_{2J}$. So by Theorem 2.1, $\beta_{II}(\boldsymbol{\mu}) \leq \beta_{II}(\boldsymbol{\mu}^*) \leq \alpha$. Now let $\boldsymbol{\mu} \in H_0$ with $\mu_1 + \mu_2 \leq 0$. Then $\boldsymbol{\mu}$ majorizes $\boldsymbol{\mu}^* = (\bar{\mu}, \bar{\mu})'$ [where $\bar{\mu} = (\mu_1 + \mu_2)/2$] and $\beta_{II}(\boldsymbol{\mu}) \leq \beta_{II}(\boldsymbol{\mu}^*)$. If we can show that $\beta_{II}(\bar{\mu}, \bar{\mu}) \leq \alpha$ for all $\bar{\mu} \leq 0$, then we will have verified that Test II is a size-α test. Furthermore, we actually need only verify that $\beta_{II}(\bar{\mu}, \bar{\mu}) \leq \alpha$ for $c_M \leq \bar{\mu} \leq 0$, because for every $\mathbf{x} \in R_1 \cup \cdots \cup R_M$, $x_1 + x_2 \geq 2c_M$. Thus by translating the problem so that (c_M, c_M) is the origin, we can use Theorem 2.2 to show that $\beta_{II}(\bar{\mu}, \bar{\mu}) \leq \beta_{II}(c_M, c_M)$ for all $\bar{\mu} < c_M$. For $\alpha = .10$, .05, and .01, we calculated $\beta_{II}(\bar{\mu}, \bar{\mu})$ for $c_M \leq \bar{\mu} \leq 0$ on a grid with spacing of .001 to find the maximum M for which $\beta_{II}(\bar{\mu}, \bar{\mu}) \leq \alpha$ for all such $\bar{\mu}$. The results are in Table 2. Test II with M equal to the tabled value is a size-α test. In the table we also list the value of $\bar{\mu}$ $(c_M \leq \bar{\mu} \leq 0)$ at which $\beta_{II}(\bar{\mu}, \bar{\mu})$ is maximized and the maximum value of $\beta_{II}(\bar{\mu}, \bar{\mu})$. But the size of Test II is α, not the value listed as $\beta_{II}(\bar{\mu}, \bar{\mu})$. The $\sup_{\boldsymbol{\mu} \in H_0} \beta_{II}(\boldsymbol{\mu})$ occurs, as with Test I and the LRT, in the limit of parameter points $(0, \mu)$ as $\mu \to \infty$.

Values of the power function of Test II for $\alpha = .10$ are given in Table 1. For $\boldsymbol{\mu}$ near $\mathbf{0}$, $\beta_{II}(\boldsymbol{\mu})$ is 1.8 times bigger than $\beta_I(\boldsymbol{\mu})$ and 9 times bigger than $\beta_L(\boldsymbol{\mu})$. In the top part of Table 1, one can see that Test II is much more nearly an unbiased test than either of the other two. But, as mentioned earlier, despite these superior power properties, Test II is probably only of theoretical interest.

Table 2. Value of M That Gives Size α for Test II

α	M	$\bar{\mu}$ at which maximum occurs	$\beta_{II}(\bar{\mu}, \bar{\mu})$
.10	9	.000	.09000
.05	19	−.884	.04906
.01	95	−.901	.00985

5. UNKNOWN VARIANCE EXAMPLE

The previous sections all dealt with models in which Σ is known. Sasabuchi (1980, 1988a,b) considered two models in which Σ was unknown. He showed that the LRT's for these models were very similar to the known-Σ LRT. The test statistics Z_i were the same, except Σ was replaced by an estimate, and z_α was replaced with a t-distribution percentile, t_α.

Because of the similarities it is natural to ask whether making the same changes in Test I will yield a test that is of size α and uniformly more powerful than the LRT. The answer is that, in general, this does not yield a size-α test. The following example illustrates this.

Consider again testing H_0: $\mu_1 \le 0$ or $\mu_2 \le 0$ versus H_1: $\mu_1 > 0$ and $\mu_2 > 0$. Let X_1 and X_2 be independent with $X_i \sim N_1(\mu_i, \sigma^2)$. Let S^2 be an independent estimate of σ^2 such that $\nu S^2/\sigma^2$ has a chi-squared distribution with ν degrees of freedom (df). Typically S^2 will be a pooled estimate of σ^2. The LRT rejects H_0 if $X_1/S > t_\alpha$ and $X_2/S > t_\alpha$, where t_α is the upper 100α percentile of a t distribution with ν df. Define $c_0, \ldots, c_J$ as in Definition 2.1 except with $t_{j\alpha}$ replacing $z_{j\alpha}$. The analog of Test I rejects H_0 if $c_j \le x_1/s \le c_{j-1}$ and $c_j \le x_2/s \le c_{j-1}$ for some $j = 1, \ldots, J$. If $h_\sigma(s)$ is the density of S, the power function of this test is

$$\beta_{\mathrm{I}}(\boldsymbol{\mu}, \sigma) = \int_0^\infty \sum_{j=1}^J P_{\boldsymbol{\mu},\sigma}(c_j s \le X_1 \le c_{j-1}s, \\ c_j s \le X_2 \le c_{j-1}s) h_\sigma(s)\,ds. \quad (5.1)$$

Using Theorem 2.2 on the integrand in (5.1), it can easily be shown that the size of this test is $\sup_{0\le\mu<\infty}\beta_{\mathrm{I}}((\mu, 0), 1)$.

We calculated $\beta_{\mathrm{I}}((\mu, 0), 1)$ for values of μ between 0 and 20 by increments of .1 using numeric integration. We did the calculations for $\alpha = .10$ and .05 and various df. The maximum value found (approximately the size of the test) is given in Table 3. In every case $\beta_{\mathrm{I}}((\mu, 0), 1)$ increased to a maximum that was greater than α and then decreased to α. So this construction does not yield a size-α test. But the size of the test does approach α as the df becomes large. For moderate or large df, this test might be preferable to the LRT, since its size is approximately α and it has higher power.

6. A UNIFORMLY MORE POWERFUL TEST IN A TWO-SIDED PROBLEM

In this section we return to the known covariance model and consider a two-sided problem involving linear inequalities. A two-sided version of the testing problem (1.1) is obtained if the alternative hypothesis is $H_1 \cup (-H_1)$, where H_1 is the set defined in (1.1). That is, consider testing

$$H_0^2\text{: } \mathbf{b}_i'\boldsymbol{\mu} \le 0 \text{ for some } i = 1, \ldots, k \text{ and} \\ \mathbf{b}_i'\boldsymbol{\mu} \ge 0 \text{ for some } i = 1, \ldots, k$$

versus

$$H_1^2\text{: } \mathbf{b}_i'\boldsymbol{\mu} > 0 \text{ for all } i = 1, \ldots, k \text{ or} \\ \mathbf{b}_i'\boldsymbol{\mu} < 0 \text{ for all } i = 1, \ldots, k. \quad (6.1)$$

Sasabuchi (1980) showed that the LRT rejects H_0^2 if $Z_i = \mathbf{b}_i'\mathbf{X}/(\mathbf{b}_i'\Sigma\mathbf{b}_i)^{1/2} \ge c$ for all $i = 1, \ldots, k$ or $Z_i \le -c$ for all $i = 1, \ldots, k$. Sasabuchi showed that, in some cases $c = z_\alpha$ gives a size-α test. We will show for a sign-testing problem that Test III, the test that rejects H_0^2 if $\mathbf{X} \in R_1 \cup \cdots \cup R_{2J}$, is a size-$\alpha$ test that is uniformly more powerful than the LRT. (The sets R_j are still the sets in Def. 2.2.) If $c = z_\alpha$, then $R_1 \cup R_{2J}$ is the rejection region for the LRT. Thus Test III is obviously a more powerful test than the LRT. The difficulty is in showing that it is a size-α test.

As mentioned in Section 1, Sasabuchi (1980) actually considered the null hypothesis that $\boldsymbol{\mu}$ was on the boundary of H_1^2. For Sasabuchi's null hypothesis, Theorem 2.1 shows that $c = z_\alpha$ is the constant that gives a size-α LRT in a broader class of problems than found by Sasabuchi in his theorems 4.1, 4.2, and 4.3. It also shows that, for this broader class, Test III is a size-α test that is uniformly more powerful than the LRT.

For the rest of this section we consider this sign-testing problem. Let $X_1, \ldots, X_p$ be independent, $X_i \sim N_1(\mu_i, \sigma_i^2)$, and consider testing

$$H_0^S\text{: } \mu_i \le 0 \text{ for some } i = 1, \ldots, p \text{ and} \\ \mu_i \ge 0 \text{ for some } i = 1, \ldots, p$$

versus

$$H_1^S\text{: } \mu_i > 0 \text{ for all } i = 1, \ldots, p \text{ or} \\ \mu_i < 0 \text{ for all } i = 1, \ldots, p. \quad (6.2)$$

The LRT rejects H_0^S if $X_i/\sigma_i \ge z_\alpha$ for all $i = 1, \ldots, p$ or $X_i/\sigma_i \le -z_\alpha$ for all $i = 1, \ldots, p$. Test III rejects H_0^S if for some $j = 1, \ldots, 2J$, $c_j \le X_i/\sigma_i \le c_{j-1}$ for all $i = 1, \ldots, p$.

To see that Test III is a size-α test, we will use majorization concepts. Let $Z_i = X_i/\sigma_i$. Then $\mathbf{Z} = (Z_1, \ldots, Z_p)' \sim N_p(\boldsymbol{\theta}, \mathbf{I}_p)$, where $\theta_i = \mu_i/\sigma_i$. For any $\boldsymbol{\mu} \in H_0^S$ there is the corresponding $\boldsymbol{\theta}$, and this $\boldsymbol{\theta}$ majorizes a vector $\boldsymbol{\theta}^*$ that has at least one coordinate equal to 0. Furthermore, the density of $\mathbf{Z}$ is Schur concave and $\bigcup_{j=1}^{2J} \{\mathbf{z}\text{: } c_j \le z_i \le c_{j-1}, i = 1, \ldots, p\}$ is a Schur-convex set. Thus we have

$$P_{\boldsymbol{\mu}}(\mathbf{X} \in \bigcup_{j=1}^{2J} R_j) \\ = P_{\boldsymbol{\theta}}(\bigcup_{j=1}^{2J} \{c_j \le Z_i \le c_{j-1} \text{ for all } i = 1, \ldots, p\}) \\ \le P_{\boldsymbol{\theta}^*}(\bigcup_{j=1}^{2J} \{c_j \le Z_i \le c_{j-1} \text{ for all } i = 1, \ldots, p\}) \\ \le \alpha,$$

Table 3. Size of Test I for Unknown Variance: Bivariate Sign-Testing Problem

	Degrees of freedom						
α	2	6	10	20	50	120	∞
.10	.1235	.1059	.1028	.1009	.1003	.1000	.1000
.05	.0702	.0564	.0535	.0514	.0505	.0501	.0500

where the first inequality is from theorem 2.1 of Marshall and Olkin (1974) and the second is from Theorem 2.1 here.

An application in which the two-sided hypotheses are of interest was described by Gail and Simon (1985). Let $\mu_i = \nu_{2i} - \nu_{1i}$, where ν_{ji} is the average response of the ith patient subset ($i = 1, \ldots, p$) to the jth treatment ($j = 1, 2$). If $\mu_i > 0$ for all $i = 1, \ldots, p$, then Treatment 2 is better in all subsets. If $\mu_i < 0$ for all $i = 1, \ldots, p$, then Treatment 1 is better in all subsets. Thus H_1^S states that the same treatment is better for all subsets. In the terminology of Gail and Simon, there is no qualitative interaction between treatment effects and patient subsets.

Gail and Simon had H_1^S: "no qualitative interaction" as the null hypothesis. So the LRT they studied is different from the one we have considered, and our results are not generally applicable in their problem. But in one case, that of $p = 2$ patient subsets, our Test III provides a uniformly more powerful size-α test in the Gail and Simon problem. To see this, let $\mu_1 = \nu_{21} - \nu_{11}$, as before, but let $\mu_2 = \nu_{12} - \nu_{22}$. Now, H_1^S: $\mu_1 > 0$ and $\mu_2 > 0$ or $\mu_1 < 0$ and $\mu_2 < 0$ states that there is a qualitative interaction, as in the Gail and Simon formulation. For this case, Zelterman (1987) constructed an approximate test that is uniformly more powerful than the Gail and Simon LRT and locally most powerful at $\boldsymbol{\mu} = \mathbf{0}$.

We have only shown that Test III is a size-α test for the special sign-testing problem (6.2). For the more general problem (6.1), Theorem 2.1 would still be useful. But the rejection region (even after transformation) might not be a Schur-convex set. Thus other techniques may be needed to find uniformly more powerful tests in the general two-sided problem.

[*Received January 1988. Revised July 1988.*]

REFERENCES

Berger, R. L. (1982), "Multiparameter Hypothesis Testing and Acceptance Sampling," *Technometrics*, 24, 295–300.

Berger, R. L., and Sinclair, D. (1984), "Testing Hypotheses Conerning Unions of Linear Subspaces," *Journal of the American Statistical Association*, 79, 158–163.

Cohen, A., Gatsonis, C., and Marden, J. I. (1983a), "Hypothesis Testing for Marginal Probabilities in a 2 × 2 × 2 Contingency Table With Conditional Independence," *Journal of the American Statistical Association*, 78, 920–929.

——— (1983b), "Hypothesis Tests and Optimality Properties in Discrete Multivariate Analysis," in *Studies in Econometrics, Time Series, and Multivariate Statistics*, New York: Academic Press, pp. 379–405.

Gail, M., and Simon, R. (1985), "Testing for Qualitative Interactions Between Treatment Effects and Patient Subsets," *Biometrics*, 41, 361–372.

Gutmann, S. (1987), "Tests Uniformly More Powerful Than Uniformly Most Powerful Monotone Tests," *Journal of Statistical Planning and Inference*, 17, 279–292.

Lehmann, E. L. (1952), "Testing Multiparameter Hypotheses," *The Annals of Mathematical Statistics*, 23, 541–552.

Marshall, A. W., and Olkin, I. (1974), "Majorization in Multivariate Distributions," *The Annals of Statistics*, 2, 1189–1200.

Nomakuchi, K., and Sakata, T. (1987), "A Note on Testing Two Dimensional Normal Mean," *Annals of the Institute of Statistical Mathematics*, 39, 489–495.

Robertson, T., and Wegman, E. J. (1978), "Likelihood Ratio Tests for Order Restrictions in Exponential Families," *The Annals of Statistics*, 6, 485–505.

Sasabuchi, S. (1980), "A Test of a Multivariate Normal Mean With Composite Hypotheses Determined by Linear Inequalities," *Biometrika*, 67, 429–439.

——— (1988a), "A Multivariate One-Sided Test With Composite Hypotheses When the Covariance Matrix Is Completely Unknown," *Memoirs of Faculty of Science, Kyushu University*, Ser. A, 42, 37–46.

——— (1988b), "A Multivariate Test With Composite Hypotheses Determined by Linear Inequalities When the Covariance Matrix Has an Unknown Scale Factor," *Memoirs of Faculty of Science, Kyushu University*, Ser. A, 42, 9–19.

Warrack, G., and Robertson, T. (1984), "A Likelihood Ratio Test Regarding Two Nested but Oblique Order-Restricted Hypotheses," *Journal of the American Statistical Association*, 79, 881–886.

Zelterman, D. (1987), "Test for Qualitative Interactions," technical report, State University of New York at Albany, Dept. of Mathematics and Statistics.